Thermophysical Affine Invariants from IR Imagery for Object Recognition*

N. Nandhakumar†
nandhu@Virginia.EDU
V. Velten‡
vvelten@mbvlab.wpafb.af.mil
Jonathan Michel†
michel@Virginia.EDU
† Dept of Electrical Engineering,
Univ. of Virginia, Charlottesville, VA 22903
‡ Wright Laboratory - WL/AARA,
Wright Patterson Air Force Base, OH 45433-7001

Submitted to: ICCV Physics-Based Modeling Workshop

Abstract

An important issue in developing a Model-Based Vision approach is the specification of features that are - (a) invariant to viewing and scene conditions, and also - (b) specific, i.e., the feature must have different values for different classes of objects. We formulate a new approach for establishing invariant features. Our approach is unique in the field since it considers not just surface reflection and surface geometry in the specification of invariant features, but it also takes into account internal object composition and state which affect images sensed in the non-visible spectrum. A new type of invariance called Thermophysical Invariance is defined. Features are defined such that they are functions of only the thermophysical properties of the imaged objects. The approach is based on a physics-based model that is derived from the principle of the conservation of energy applied at the surface of the imaged object.

Keywords: Object Recognition, Feature Invariance

*This research was supported by the AFOSR contract F49620-93-C-0063, the AFOSR grant LRIR-93WL001, an AFOSR Laboratory Graduate Fellowship, ARPA contract F33615-94-C-1529 and the National Science Foundation Contract IRI-91109584.

1. Introduction

Object recognition requires robust and stable features that are separable in feature space. An important characteristic of these features is that they be invariant to scene conditions, such as illumination, and changes in viewpoint/object pose. The formulation of invariant features, and the quantitative analysis of feature variance is currently being addressed by a number of researchers in the computer vision community, and has led to the establishment of a currently incomplete but growing theory of feature invariance. Such efforts have primarily considered reflected-light imagery – formed by sensing visible wavelength energy. This investigation has resulted in a number of distinct (yet related) approaches that may be loosely grouped into three categories: (1) Geometric Invariance (GI), (2) Quasi-Invariance (QI), and (3) Intensity-based Invariance (II).

Geometric invariants have been investigated since the inception of the field of image analysis in the early 1960s (actually, such investigation can be traced back to the onset of photogrammetry in the 19th century). The main issue here is the investigation of features that are invariant with respect to changes in viewpoint. Geometric invariants come in two basic "flavors", algebraic and differential. Algebraic invariants are based on the global configuration of features extracted from an object, and involve the notion of algebraic shapes, e.g., shapes are analytically expressed as 2D conics, and invariant relationships are identified between conics belonging to an object. Differential invariants are polynomial expressions involving the local curvature properties of 2D and 3D curves and are computed for each point on the curve. Thus differential invariants are actually parameter space "signatures" (e.g., a locus of point in an abstract parameter space) that are unique to the object, so that differences between two parameter space signatures define different objects. A close relationship (an equivalence in some cases) has been shown between some formulations of differential invariants and algebraic invariants [Forsyth et al., 1991]. There are several examples of geometric (actually algebraic) invariants of planar configurations under projective transformation, such as the cross ratio

using 4 collinear points, 5 coplanar points with no three being collinear, a conic and two non-tangent lines, and a pair of conics.

Quasi-Invariance (QI) can be thought of as a relaxation of the central notion of GI [Binford et al., 1989, Zerroug and Nevatia, 1993]. A Quasi-Invariant is a property of a geometric configuration that is almost invariant under a class of imaging transformations. Formally, a geometric configuration is a QI if the linear term(s) in the Taylor series expansion of the configuration with respect to the parameters of the imaging transformation vanish. This has the effect of making the QI measure nearly constant over a large region of the viewing hemisphere, and rapidly diverging only as the angle between the view direction and the surface normal approaches large values. This behavior lends itself to probabilistic modeling and the use of reasoning schemes such as Bayesian evidence accrual. A detailed study of the variation of "invariant" features with viewpoint has been undertaken [Burns et al., 1993].

Intensity-based Invariants (II's) are functions of image intensities that yield values which are invariant to scene illumination and viewpoint. To date, some investigation of II's has been conducted for visible imagery, but practically none has been reported for non-visible imagery. Examples of II's in computer vision include color features for object recognition [Klinker et al., 1988, Healey, 1991, Healey and Slater, 1994] and polarization cues for material identification [Wolf, 1990]. A more direct example of this approach computes ratios of albedos of homogeneous image intensity patches within objects in visible imagery [Nayar and Bolle, 1993].

Non-visible modalities of sensing have been shown to greatly increase the amount of information that can be used for object recognition. A very popular and increasingly affordable sensor modality is thermal imaging - where non-visible radiation is sensed in the long-wave infrared (LWIR) spectrum of $8\mu m$ to $14\mu m$. The current generation of LWIR sensors produce images of contrast and resolution that compare favorably with broadcast television quality visible light imagery. However, the images are

no longer functions of only surface reflectance. As the wavelength of the sensor transducer passband increases, emissive effects begin to emerge as the dominant mode of electromagnetic energy exitance from object surfaces. The (primarily) emitted radiosity of LWIR energy has a strong dependence on internal composition, properties, and state of the object such as specific heat, density, volume, heat generation rate of internal sources, etc. This dependence may be exploited by specifying image-derived invariants that vary only if these parameters of the physical properties vary.

In this paper we describe the use of the principle of conservation of energy at the surface of the imaged object to specify a functional relationship between the object's thermophysical properties (e.g., thermal conductivity, thermal capacitance, emissivity, etc.), scene parameters (e.g., wind temperature, wind speed, solar insolation), and the sensed LWIR image gray level. We use this functional form to derive invariant features that remain constant despite changes in scene parameters/driving conditions. In this formulation the internal thermophysical properties play a role that is analogous to the role of parameters of the conics, lines and/or points that are used for specifying geometric invariants when analyzing visible wavelength imagery. Thus, in addition to the currently available techniques of formulating features that depend only on external shape and surface reflectance discontinuities, the phenomenology of LWIR image generation can be used to establish new features that "uncover" the composition and thermal state of the object, and which do not depend on surface reflectance characteristics.

The derivation of thermophysical invariants (TI's) from non-visible wavelength imagery, the evaluation of the performance of these invariants, and their use in object recognition systems poses several advantages. The main advantage of this approach is the potential availability of a number of new (functionally independent) invariants that depend on internal compositional properties of the imaged objects. Note that it is possible to evaluate the behavior of thermophysical invariants using ground truth data consisting of images of objects of known composition and internal state. This additional

information can be used to augment/complement the behavior of GI's. One way in which GI's can be integrated with TI's for object recognition is as follows: (1) Parametric curves and/or lines are extracted from an LWIR image. (2) The curves are used to compute GI's which are in turn used to hypothesize object identity and pose, and (3) TI's are computed for this hypothesis and compared to a stored model library for verification. Some details of this approach are presented later.

The ideas presented in this paper are continuations/extensions of previous and ongoing research in thermophysical model-based interpretation of LWIR imagery. A brief description of this thermophysical approach is presented in section 2, followed by the formulation of a method to derive themophysical invariants in section 3. Preliminary experimental results of applying this new approach to real imagery are presented in section 4, which is followed by a disussion of the behavior of the new method, issues to be considered in using this method for object recognition, and issues that remain to be explored.

2. A Thermophysical Approach to IR Image Analysis

An intuitive approach to thermo-physical interpretation of LWIR imagery is given in [Gauder, et al, 1993]. This approach rests upon the following observation, termed the "Thermal History Consistentcy Constraint" and analogous to Lowe's well known Viewpoint Consistentcy Constraint [Lowe, 1987]: "The temperature of all target features for a passive target must be consistent with the heat flux transfer resulting from exposure to the same thermal history." In [Gauder, et al, 1993] this constraint is exploited by analyzing objects to locate components that are similar in terms of thermo-physical properties and then examining a temporal sequence of calibrated LWIR data to experimentally assess the degree to which such thermo-physically similar components exhibit similar temperature state temporal behavior. Such analysis was shown to lead to formulation of simple intensity ratio features exhibiting a strong degree of temporal stability that could be exploited provided: (1) thermally homogenous regions in the LWIR image corresponding to the thermo-physically similar object com-

Figure 1: Energy exchange at the surface of the imaged object. Incident energy is primarily in the visible spectrum. Surfaces loses energy by convection to air, via radiation to the amosphere, and via conduction to the interior of the object.

ponents could be reliably segmented, and (2) a target-specific geometric reference frame is available in order to correctly associate extracted regions with candidate object components.

To avoid the difficulties inherent in assumptions (1) and (2) above an alternative technique applicable to overall object signatures was suggested in [Gauder, et al, 1993]. An analysis of typical LWIR 'lumped parameter' object temperature modeling approaches suggests that over small time scales object temperature can be crudely modeled by a small dimensional linear system with algebraically seperable spatial and temporal components. Ratios of spatial integrals of temperature with a simple set of orthronormal 2D polynomials (obtained from applying Gramm-Schmidt to $1, x, y, xy, x^2y, xy^2, x^3$ and y^3) were tried. Some of the resulting functions were nearly constant with time when measured against 24 hours of LWIR imagery of a complex object (a tank) but no experimentation was done with multiple objects to examine between and within-class seperation, so little can be drawn in the way of a substantive conclusion with respect to utility as an object identification technique.

A physics-based approach that attempts to establish invariant features which depend only on thermophysical object properties is reported in [Nandhakumar and Aggarwal, 1988, Nandhakumar,

1990]. A thermophysical model is formulated to allow integrated analysis of thermal and visual imagery of outdoor scenes. An overview of this approach, it's advantages and limitations is presented below, and an improvement to this thermophysical approach using algebraic invariance theory is described in section 3.

At the surface of the imaged object (figure 1) energy absorbed by the surface equals the energy lost to the environment.

$$W_{abs} = W_{lost} \tag{1}$$

Energy absorbed by the surface is is given by

$$W_{abs} \;=\; W_I \; cos\theta_I \; \alpha_s \;, \tag{2}$$

where, W_I is the incident solar irradiation on a horizontal surface and is given by available empirical models (based on time, date and latitude of the scene) or by measurement with a pyranometer, θ_i is the angle between the direction of irradiation and the surface normal, and α_s is the surface absorptivity which is related to the visual reflectance ρ_s by $\alpha_s = 1 - \rho_s$. Note that it is reasonable to use the visual reflectance to estimate the energy absorbed by the surface since approximately 90% of the energy in solar irradiation lies in the visible wavelengths [Incropera and DeWitt, 1981].

A simplified shape-from-shading approach was used to compute $cos\theta_I$ and α_s from the visual image. The surface temperature was estimated from the thermal image based on an appropriate model of radiation energy exchange between the surface and the infrared camera.

The energy lost by the surface to the environment was given by

$$W_{lost} \;=\; W_{cd} \;+\; W_{cv} \;+\; W_{rad}, \tag{3}$$

where, W_{cv} denotes the heat convected from the surface to the air which has temperature T_{amb} and velocity V, W_{rad} is the heat lost by the surface to the environment via radiation and W_{cd} denotes the

Figure 2: Equivalent thermal circuit of the imaged object. Here, internal heat sources are assumed to be non-existent. The conducted energy is used to charge/discharge the lumped thermal capacitance of the object.

heat conducted from the surface into the interior of the object. The radiation heat loss is computed from:

$$W_{rad} \;=\; \epsilon\ \sigma\ (T_s^4 - T_{amb}^4), \tag{4}$$

where, σ denotes the Stefan-Boltzman constant, T_s is the surface temperature of the imaged object, and T_{amb} is the ambient temperature.

The convected heat transfer is given by

$$W_{cv} \;=\; h(T_s \;-\; T_{amb}) \tag{5}$$

where, h is the average convected heat transfer coefficient for the imaged surface, which depends on the wind speed, thermophysical properties of the air, and surface geometry [Incropera and DeWitt, 1981].

Considering a unit area on the surface of the imaged object, the equivalent thermal circuit for the surface is shown in figure 2. C_T is the lumped thermal capacitance of the object and is given by

$$C_T = DVc$$

where, D is the density of the object, V is the volume, and c is the specific heat. The resistances are given by:

$$R_{cv} = \frac{1}{h} \qquad \text{and} \qquad R_{rad} = \frac{1}{\epsilon\sigma(T_s^2 + T_{amb}^2)(T_s + T_{amb})} \tag{6}$$

It is important to note that implicit in the above formulation is that a low Biot number was assumed, viz., the surface was considered to be a thin plate.

2.1. Analyzing A Single Data Set

It is clear from figure 2 that the conduction heat flux W_{cd} depends on the lumped thermal capacitance C_T of the object. A relatively high value for C_T implies that the object is able to sink or source relatively large amounts of heat. An estimate of W_{cd}, therefore, provides us with a relative estimate of the thermal capacitance of the object, albeit a very approximate one. W_{cd} is useful in estimating the object's ability to sink/source heat radiation, a feature shown to be useful in discriminating between different classes of objects.

In order to minimize the feature's dependence on differences in absorbed heat flux, a normalized feature was defined to be the ratio $R = W_{cd}/W_{abs}$. The values are lowest for vehicles, highest for vegetation and in between for buildings and pavements. Classification of objects using this property value is discussed in [Nandhakumar and Aggarwal, 1988b].

2.2. Analyzing A Temporal Data Sequence

The availability of temporal sequences of registered thermal and visual images makes possible a direct estimation of C_T and hence a more reliable estimate of the imaged object's relative ability to sink/source heat radiation. Since

$$W_{cd} = C_T \frac{dT_s}{dt} \tag{7}$$

a finite (backward) difference approximation to this equation may be used for estimating C_T as

$$C_T = W_{cd} \frac{(t_2 - t_1)}{(T_s(t_2) - T_s(t_1))} \tag{8}$$

where, t_1 and t_2 are the time instants at which the data were acquired, $T_s(t_1)$ and $T_s(t_2)$ are the corresponding surface temperatures, and W_{cd} is the conducted heat flux which is assumed to be

constant during the time interval.

2.3. Limitations of The Previous Approach

The above approach is powerful in that it makes available features that are completely defined by internal object properties, and hence may be considered to be deterministic. Hence, feature values may be compared with accurate *ground truth* values computed from known physical properties of test objects – one of the major advantages of using physics-based/phenomenological models as compared to statistical models.

There are several factors that limit the performance of the above approach. These are listed below and a method to minimize these limitations is addressed in the next section.

Systemic Errors in Sensor Fusion.

The thermal and visual image pairs may not be perfectly registered. Also, segmentation errors typically cause a large portion of an object to be included with small portions of a different object in one region. These errors give rise to meaningless values of the surface energy estimates at/near the region boundaries. However, stable estimates are available in the interior of each of the regions. Also, shape from shading techniques perform poorly unless images are relatively noise-free and have high resolution [Zhang et al., 1994].

These errors result in a significant number of inaccurate estimates of the surface energy exchange components. The histogram of values of the ratio R tends to be heavy-tailed and skewed. A statistically robust scheme for computing the thermophysical feature R was proposed to minimize this drawback [Nandhakumar, 1994]. However, the computational complexity for such a technique is very high, and the following drawbacks below were not adequately overcome.

Unknown Surface/Scene Parameters.

Direct application of the energy exchange model as described above requires *a priori* knowledge of

Figure 3: The equivalent thermal circuit for the extended model that separates the stored energy component and the conduction component to the interior of the object.

several surface and scene parameters such as emissivity, wind speed, solar insolation, etc. Values for these were either assumed, hypothesized, or empirically obtained. It would be beneficial to formulate an approach where knowledge of these parameters is not imperative.

Variation in R.

The value of the thermophysical feature, R, was found to be only weakly invariant. While separation between classes was preserved, the range of values of R for each class was observed to vary with time of day and season of year. Also, the feature R was able to only separate very broad categories of objects, such as automobiles, buildings, and vegetation - but lacked the specificity to differentiate between different models of vehicles.

An improved formulation for establishing thermophysical features is described in section 3, wherein the feature is constrained to be invariant to affine transformation of the driving (scene) conditions.

3. Formulation of Thermophysical Affine Invariants (TPI's)

An improved approach for computing invariant features that depend on the internal, thermophysical properties of the imaged object, and that are invariant to driving (scene and surface) parameters is based on a reformulation of the model outlined in section 2. First, to account for reasonable values of Biot numbers, viz. for objects other than thin plates, the conducted energy component W_{cd} is decomposed into two terms - W_{st} which alters the stored, internal energy of an elemental volume

Figure 4: We need a collection of five point in 5D measurement space to compute the affine transformation, T_i, between two scenes. Two different sets of 5 points each can be used to define an absolute invariant provided $\det(T_1) = \det(T_2)$.

at the surface, and W_{cnd} which is conducted to the interior of the object. Note that W_{st} is given by eqn (7), while $W_{cnd} = -k\ dT/dx$, where k is the thermal conductivity of the material, and x is distance below the surface. Here, we assume that lateral energy conduction is insignificant compared to conduction along the direction normal to the surface. Figure 3 shows the equivalent thermal circuit for this extended model.

The energy balance equation, $W_{abs} = W_{rad} + W_{cv} + W_{st} + W_{cnd}$ may be rewritten in the following linear form:

$$a^1 x_1 + a^2 x_2 + a^3 x_3 + a^4 x_4 + a^5 x_5 = 0. \tag{9}$$

Using the expressions for the various energy components as presented in section 2 we can express each

term in the above expression as:

$$\begin{aligned} a^1 &= cos\theta_I & x_1 &= W_I\alpha_s \\ a^2 &= -\sigma\ (T_s^4 - T_{amb}^4) & x_2 &= \epsilon \\ a^3 &= -(T_s - T_{amb}) & x_3 &= h \\ a^4 &= C_T & x_4 &= -\frac{dT_s}{dt} \\ a^5 &= kT_s & x_5 &= 1 - \frac{T_{int}}{T_s} \end{aligned} \tag{10}$$

The term $a^5 x_5$ denotes W_{cnd} expressed in a finite difference form.

Note that a calibrated LWIR image provides radiometric temperature. Hence a^2 and a^3 can be computed from the LWIR image alone (and knowledge of the ambient temperature), while a^1 and a^4 are known when the identity and pose of the object is hypothesized, and a^5 is computed using the image and the hypothesis. The "driving conditions", or unknown scene parameters[1] that change from scene to scene are given by the x_i. For each pixel in the thermal image eqn (10) defines a hyperplane in 5-D space.

Consider two different LWIR images of a scene obtained under different scene conditions and from different viewpoints. Consider N points on the object that (a) are visible in both views, and (b) have been selected to lie on different components of the object which differ in material composition and/or surface normal direction. Assume (for the nonce) that the object pose for each view, and point correspondence between the two views are available (or hypothesized). A point in each view yields a contravariant tensor a^i as defined by eqn (10). Let the collection of these tensors be denoted by $a_k^i, k = 1, 2, \ldots, N$ for the first scene/image and $b_k^i, k = 1, 2, \ldots, N$ for the second scene. For the k-th point we denote the measurement tensor as $\mathbf{a_k}$ for the first view, and as $\mathbf{b_k}$ for the second view, and the driving conditions tensor as $\mathbf{x^k}$.

[1] Here, we use the Einstein notation to denote the image-based measurement vector as a contravariant tensor,a^i, while the changing scene conditions form a covariant tensor,x_i. Therefore, $a^i x_i = 0$.

We assume that the scene/driving conditions, $\mathbf{x^k}$, in the two scenes are related by an affine transformation. The justification of this assumption is discussed below. We have found, empirically, that this assumption holds when the points are selected using the method discussed later in this paper. Since the $\mathbf{x^k}$ are transformed affinely, then it follows that the $\mathbf{a_k}$ are also transformed affinely. Note that an affine transformation from one scene to another is trivial to obtain if we have only five points that generate five non-coplanar tensors in our 5-D measurement space. Consider one such subset of 5 of the N points, and denote them as $i, j, l, m,$ and n.

The determinant

$$d(\mathbf{a_i a_j a_l a_m a_n}) = \begin{vmatrix} a_i^1 & a_i^2 & a_i^3 & a_i^4 & a_i^5 \\ a_j^1 & a_j^2 & a_j^3 & a_j^4 & a_j^5 \\ a_l^1 & a_l^2 & a_l^3 & a_l^4 & a_l^5 \\ a_m^1 & a_m^2 & a_m^3 & a_m^4 & a_m^5 \\ a_n^1 & a_n^2 & a_n^3 & a_n^4 & a_n^5 \end{vmatrix} \tag{11}$$

defines the volume of the oriented parallelopiped formed by the pencil of the five contravariant tensors $\mathbf{a_i}, \mathbf{a_j}, \mathbf{a_l}, \mathbf{a_m}, \mathbf{a_n}$. The above determinant is a relative invariant to the affine transformation [Gurevich, 1964], i.e.,

$$d(\mathbf{a_i a_j a_l a_m a_n}) = \delta_{ijlmn} \times d(\mathbf{b_i b_j b_l b_m b_n}) \tag{12}$$

where δ_{ijlmn} is the determinant of the affine transformation, T_{ijlmn}, which relates the measurement tensors, i.e., $\mathbf{a_k} = \mathbf{b_k} T_{ijlmn}$, $k \in \{i, j, l, m, n\}$.

Consider another set of five points in which at least one point is different from the previous set (see figure 4). Denote this second set as $\{p, q, r, s, t\}$. Again, assume that the measurement tensors for this collection of points undergo an affine transformation from the first scene to the second, and denote this transformation by T_{pqrst}.

$$d(\mathbf{a_p a_q a_r a_s a_t}) = \delta_{pqrst} \times d(\mathbf{b_p b_q b_r b_s b_t}) \tag{13}$$

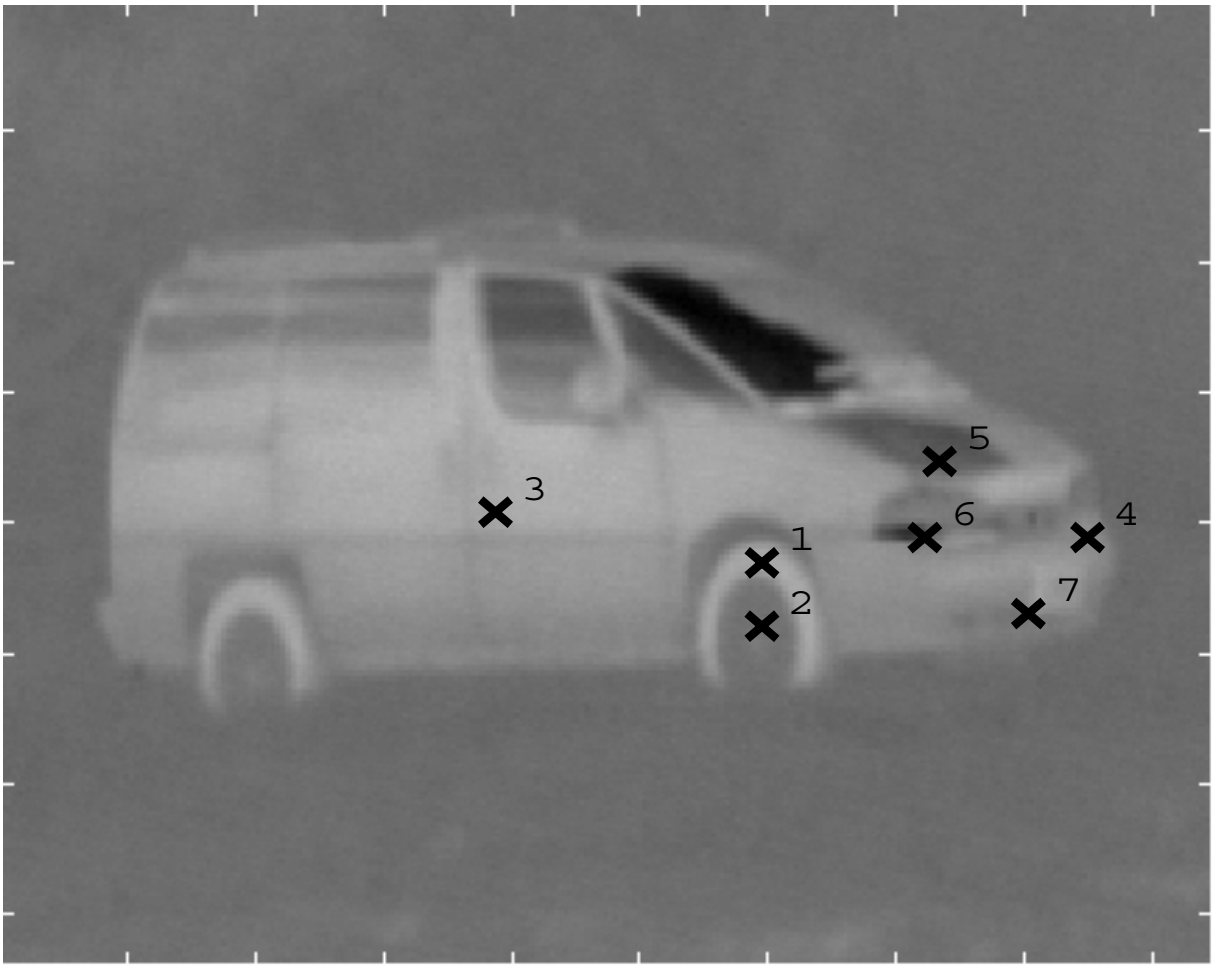

Figure 5: The van with points selected on the surface with different material properties and/or surface normals. Point 1: Vulcanized Rubber, 2: Aluminum Alloy, 3: Polystyrene, 4: Steel, 5: Polypropylene, 6: Steel, 7: Polypropylene

where $\delta_{pqrst} = \det(T_{pqrst})$. Hence, if $\delta_{ijlmn} = \delta_{pqrst}$, then we can define an absolute invariant as

$$I = \frac{d(\mathbf{a_i a_j a_l a_m a_n})}{d(\mathbf{a_p a_q a_r a_s a_t})} \tag{14}$$

Note that the existence of affine transformations T_{ijlmn} and T_{pqrst} is easy to ensure by selecting the surface points appropriately (discussed in section 5). The selection of two sets of five points such that $\delta_{ijlmn} = \delta_{pqrst}$ holds can be attempted as a data-driven training task as follows.

Calibrated LWIR imagery from different object classes are obtained at different times of day and

different seasons of the year. N points are picked on an object – on distinctive components that differ in material composition and/or surface normal. Consider the image from time t_u and the image from t_v, $u \neq v$. The measurements at t_u along with the hypothesis of the identity of the object form the tensors $\mathbf{a_k}$. Similarly, image information at time t_v is used to form the measurement tensors $\mathbf{b_k}$.

All combinations of two sets of five points each, $\{i, j, l, m, n\}$ and $\{p, q, r, s, t\}$, are examined. The measurement matrices $(\mathbf{a_i a_j a_l a_m a_n})$, $(\mathbf{b_i b_j b_l b_m b_n})$, $(\mathbf{a_p a_q a_r a_s a_t})$, and $(\mathbf{b_p b_q b_r b_s b_t})$ are constructed. The transformations T_{ijlmn} and T_{pqrst}, if they exist, and the their determinants δ_{ijlmn}, and δ_{pqrst} are computed. The two sets that best satisfy $\delta_{ijlmn} = \delta_{pqrst}$ for different choices of pairs of scenes/images, i.e. different choices of t_u and t_v, are selected. With N points, the number of possible choices of the pair of sets of points is given by

$$n_{_N} \;=\; \frac{1}{2}\;\left(\frac{N!}{(N-5)!5!}\right)\;\left(\frac{N!}{(N-5)!5!}-1\right) \tag{15}$$

For example, $n_6 = 15$, $n_8 = 1,540$, and $n_{11} = 106,491$.

In order for the invariant feature to be useful for object recognition the value of I must be different if the measurement vector is obtained from a scene that does not contained the hypothesized object and/or the hypothesized pose is incorrect. Hence, the search for the optimal sets of points conducted during training phase should also take into consideration class separability, in addition to intra-class variation.

The linear form, eqn (10), must be slightly modified for interpreting LWIR imagery acquired at night. Since solar insolation is nonexistent, W_{abs} as defined above is zero, and the energy balance model has only four terms. Hence, the measurement tensor is four dimensional, and we consider sets of four points in evaluating the absolute invariant. A separate training phase is required, resulting in the specification of a different choice of points for night-time use.

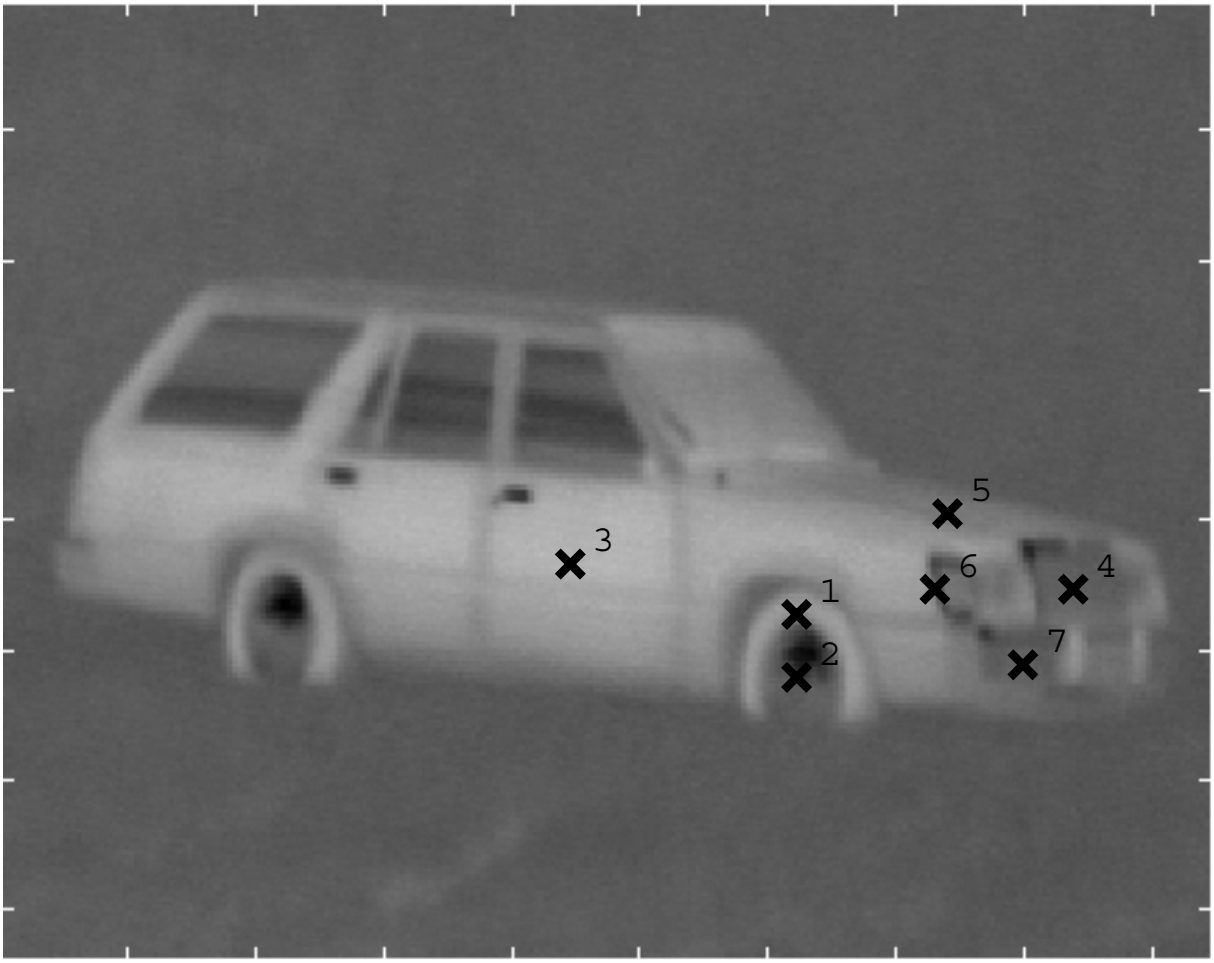

Figure 6: The car with points selected on the surface with different material properties and/or surface normals. Point 1: Steel, 2: Vulcanized Rubber, 3: Steel, 4: Chromed Steel, 5: Steel, 6: Polycarbonate, 7: Carbon Steel

4. Experimental Results

The method of computing thermophysical affine invariants discussed above was applied to real LWIR imagery acquired at different times of the day. Two types of vehicles were imaged: A van and a car (figures 6 and 5). Several points were selected (as indicated in the figures) on the surfaces of different materials and/or orientation. The measurement tensor given by eqn (10) was computed for each point, for each image/scene.

The method used to select optimal sets of points $\{i, j, l, m, n\}$ and $\{p, q, r, s, t\}$ was similar to that described in section 3 – however, instead of using the equivalence of the determinants of the two

Time	Set A-1	Set A-2
9 am	-0.1142	0.1398
10 am	-0.1095	0.1230
11 am	-0.1098	0.1085
12 n	-0.1049	0.1186
1 pm	-0.1051	0.0919

Table 1: Variation of feature values for two different choices of points. A-1 consists of point sets $\{1,4,5,6,7\}$ and $\{2,4,5,6,7\}$. A-2 consists of point sets $\{1,3,4,5,6\}$ and $\{2,4,5,6,7\}$. The numbers correspond to the labels in the image of the Van.

Time	Set B-1	Set B-2
9 am	-0.1805	0.1238
10 am	-0.1251	0.0693
11 am	-0.0846	0.0898
12 n	-0.1324	0.0829
1 pm	-0.0602	0.1971

Table 2: Variation of feature values for two different choices of points. B-1 consists of point sets $\{1,2,3,4,6\}$ and $\{1,3,5,6,7\}$. B-2 consists of point sets $\{1,2,4,6,7\}$ and $\{2,3,5,6,7\}$. The numbers correspond to the labels in the image of the Car.

affine transformations as the selection criterion, we used the variance in the values of the feature computed for different scenes (i.e., images obtained at different times of day) containing the object. Many different pairs of five-point-sets yielded features with low variance from scene to scene. Table 1 shows values of the feature for two different choices of sets of points for the van, and table 2 shows values for two different choices of points for the car.

As mentioned in section 3 one must consider inter-class behavior as well as intra-class behavior. To investigate this we adopted the following procedure. Given an image of a vehicle, (1) assume the pose of the vehicle is known, then (2) use the front and rear wheels to establish an object centered reference frame. The center of the wheel is used as the origin, and center of the front wheel is used to specify the direction and scaling of the axes. The coordinates of the selected points are expressed in terms of this 2D object centered frame. Thus when a van vehicle is hypothesized for an image actually obtained of a car or some unknown vehicle, the material properties of the van are used, but image measurements are obtained from the image of the car at locations given by transforming the

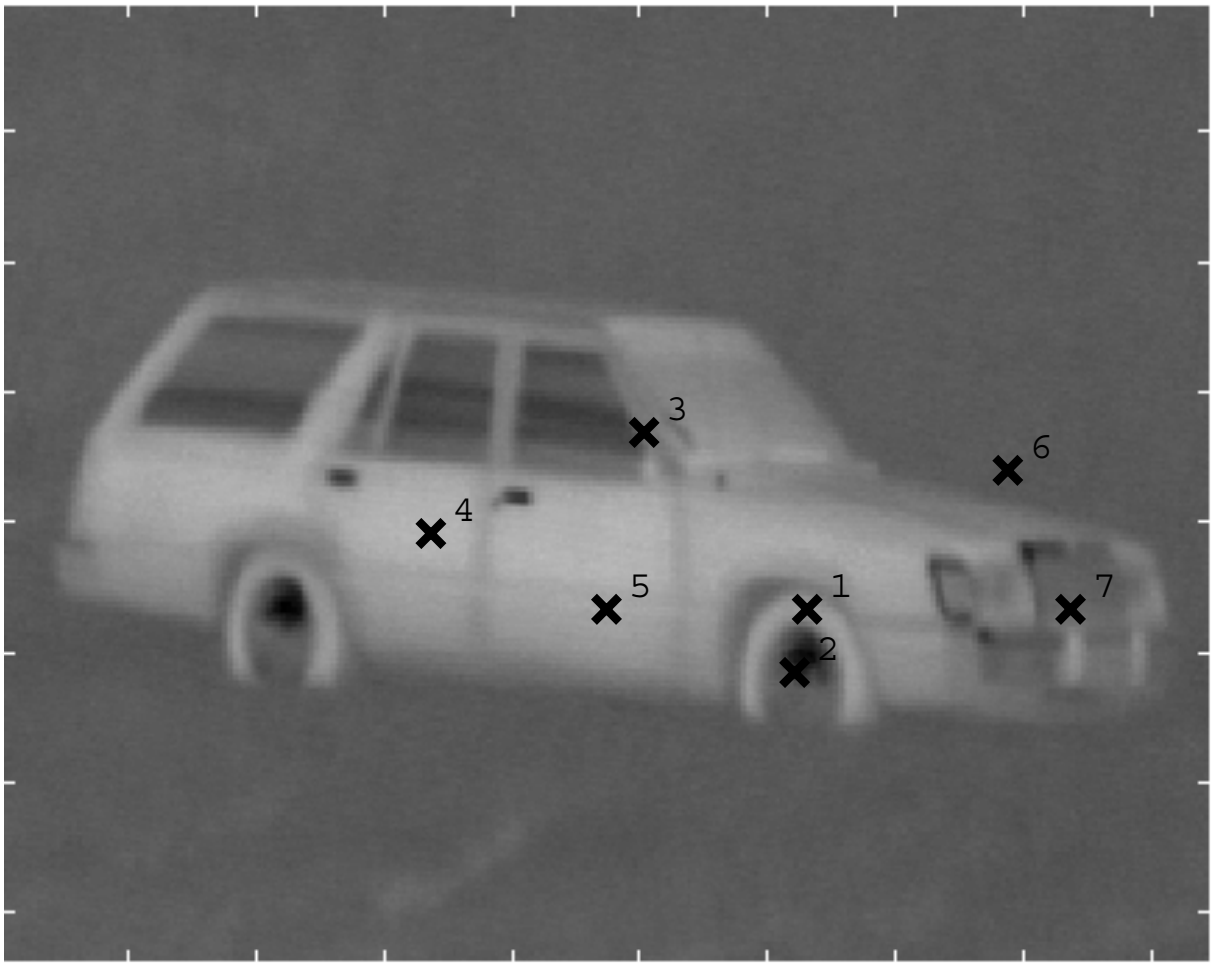

Figure 7: For the car image, if the van is hypothesized with pose as shown in figure 5 the locations of the corresponding points in the car reference frame are as shown.

coordinates of the van points (in the van center coordinate frame) to the image frame computed for the unknown vehicle.

Table 3 shows inter-class and intra-class variation when a van is hypothesized, and for images obtained at five different times in the day. Table 4 shows inter-class and intra-class variation when the car is hypothesized. Such investigation showed that the set of points B-1 produced almost identical values irrespective of the source of the measurements. This set has good (low) intra-class variation but poor inter-class separation, and it does not distinguish the car from the van. The sets of points A-1 and B-2 have good (low) intra-class variation and acceptable intra-class variation.

Time of Day	Hypothesis: Van Data from: Car	Hypothesis: Van Data from: Van
9 am	-7.6843	-0.1142
10 am	-3.9699	-0.1095
11 am	-46.2402	-0.1098
12 n	2.2820	-0.1049
1 pm	-0.8985	-0.1051

Table 3: Inter-class variation vs. intra-class variation for set A-1. Points and thermophysical properties are chosen for a van hypothesis. The middle column shows feature values computed when the measurement tensor is obtained from the car.

Time of Day	Hypothesis: Car Data from: Van	Hypothesis: Car Data from: Car
9 am	-0.3515	-0.1805
10 am	0.1434	-0.1251
11 am	-0.0627	-0.0846
12 n	-1.1355	-0.1324
1 pm	0.0296	-0.0602

Table 4: Inter-class variation vs. intra-class variation for point set B-1. Points and thermophysical properties are chosen for the car hypothesis. The middle column shows feature values computed when the measurement tensor is obtained from the van. This set exhibits poor inter-class separation.

Time of Day	Hypothesis: Car Data from: Van	Hypothesis: Car Data from: Car
9 am	-1.0216	0.1238
10 am	-1.2108	0.0693
11 am	-1.3368	0.0898
12 n	-1.1245	0.0829
1 pm	-0.0437	0.1971

Table 5: Inter-class variation vs. intra-class variation for point set B-2. Points and thermophysical properties are chosen for the car hypothesis. The middle column shows feature values computed when the measurement tensor is obtained from the van. This set of points shows good inter-class separation.

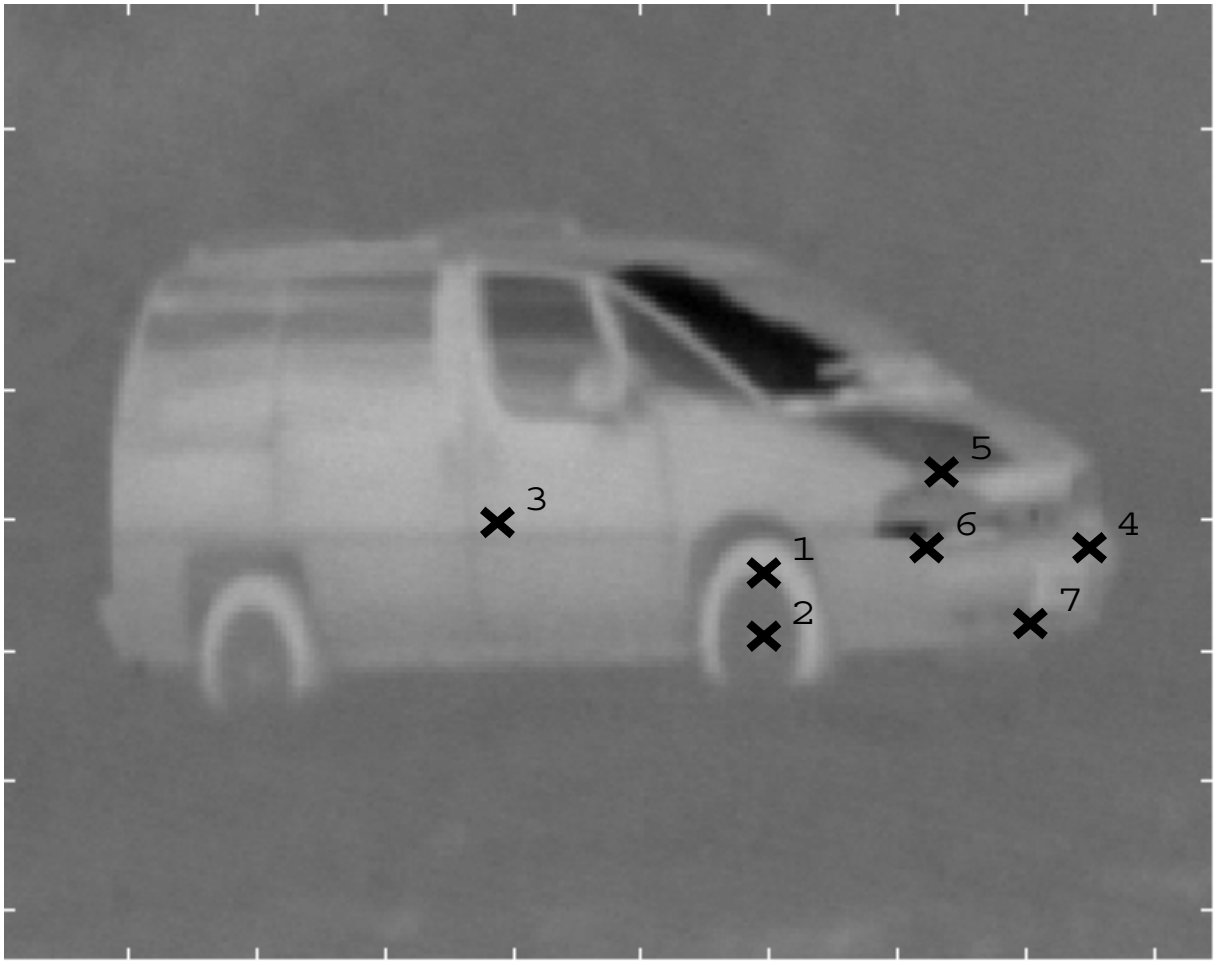

Figure 8: For the van image, if the car is hypothesized with pose as shown in figure 6 the locations of the corresponding points in the van reference frame are as shown.

5. Discussion

The approach described above is promising in that it makes available features that are (1) invariant to scene conditions, (2) able to separate different classes of objects, and (3) based on physics based models of the many phenomena that affect LWIR image generation.

The specification of optimal sets of points for high inter-class separation and low intra-class variation is a crucial task in this approach. This is a complex search problem, and it is not clear that a solution will always exist for a collection of object classes. Note that different aspects of an object may be imaged – the set of visible points differ for each aspect. The complexity of the search

task is compounded by attempting to ensure inter-class separation in the presence of erroneous pose hypothesis.

Some criteria for the choice of point sets are obvious. Points should not be chosen such that the measurement matrix has less than full rank. This occurs, for example, when the five points lie on surfaces with identical surface normals and identical thermophysical properties. When the two sets of points have four points in common, and when the remaining two points are on different parts of the imaged object, but lie on materials that are identical/similar, then these points will form an invariant that has magnitude of one. Such invariants are useful only if there exists a unique set, or limited number of sets, of points that produces this value. Other thermophysical criteria for point set selection need to be investigated.

Considering a simple case of the point set selection gives insight into the formulation of the invariant feature. Recall that two sets of five points are required to form the feature value. The simplest method to form two sets is to replace one point in the first set with a distinct point in the second set. For a given scene we may construct: $A = [\mathbf{a_1 a_2 a_3 a_4 a_5}]$ and $A' = [\mathbf{a_1 a_2 a_3 a_4 a_6}]$, the measurement matrices with measurement tensors, $\mathbf{a}$'s, as defined in section 3. Consider $\vec{n}$, which spans the null space of $A^* = [\mathbf{a_1 a_2 a_3 a_4}]$, i.e. $A^*\vec{n} = 0$. The feature value, $\frac{|A|}{|A'|}$, is non-zero if and only if $\mathbf{a}_5^T \cdot \vec{n} \neq 0$ and $\mathbf{a}_6^T \cdot \vec{n} \neq 0$. Furthermore, it can be shown that:

$$\frac{|A|}{|A'|} = \frac{\mathbf{a}_5 \cdot \vec{n}}{\mathbf{a}_6 \cdot \vec{n}} \tag{16}$$

The proof of this is omitted for brevity. A topic of future research is to $\vec{n}$ such that it varies little from scene to scene. If such a formulation can be found it will be possible to pre-compute $\vec{n}$ for a given object. That will reduce the required measurements in the image to two points.

There remain many issues that are not well understood and require further research. An important assumption/observation that requires justification from a thermophysical viewpoint is that the tensor

of driving conditions (for the two sets of points used to compute the invariant) undergoes an affine transformation from scene to scene, and that the affine transformations for the two sets have identical determinants for any two scenes. While an affine transformation can model a wide range of changes in the driving conditions, and hence is intuitively a good choice - a thermophysical interpretation of this assumption/observation would be helpful in understanding the behavior of invariants for different choice of points on the object. Another important issue is to choose a linear form that is not homogeneous - by replacing the zero in the right hand side of eqn (9) with terms from the left hand side. This would reduce the dimensionality of the measurement vector, and depending on the choice of terms – eliminate the requirements for hypothesis of pose and hypothesis of specific thermophysical parameter values, but at the risk of impaired invariance and separability.

The hypothesis of object pose and identity is best achieved by employing geometrical invariance techniques [Forsyth et al., 1991]. For example, conics may be fit to wheels which manifest high contrast in LWIR imagery, and their parameter values may be used to compute GI's. This may be employed to generate object identity and pose that may be verified by the thermophysical invariance scheme described above. Future effort will be devoted to: the integration of the above scheme with GI's to produce a complete system, the study of the nature of scene-to-scene transformation of driving conditions and justification of the affineness of this transformation, and a detailed exploration of the performance of the scheme when applied to a significant collection of objects, aspects, and scene conditions.

References

[1] T. Binford, T.S. Levitt, and W.B. Mann, "Bayesian Inference in Model-Based Vision", *Uncertainty in AI, 3*, L.N. Kanal, T.S. Levitt, and J.F. Lemmer, (Ed's), Elsevier, 1989.

[2] J.B. Burns, R.S. Weiss, and E.M. Riseman, "View Variation of Point-Set and Line-Segment Features", *IEEE Trans PAMI*, vol 15, no 1, Jan 1993.

[3] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, C. Rothwell, "Invariant Descriptors for 3D Object Recognition and Pose", *IEEE Trans PAMI*, vol 13, no 12, Oct 1991

[4] M.J. Gauder, V.J. Velten, L.A. Westerkamp, J. Mundy, and D. Forsyth, "Thermal Invariants for Infrared Target Recognition", *ATR Systems and Technology Conf*, 1993.

[5] G.B. Gurevich, "Foundations of the Theory of Algebraic Invariants", (translated by J.R.M. Radddock and A.J.M. Spencer) P. Noordhoff Ltd - Groningen, The Netherlands, 1964

[6] G. Healey, "Using Color to Segment Images of 3-D Scenes", *Proc SPIE Conf Applications of AI*, vol. 1468, 1988, Orlando, FL, pp. 814-825.

[7] G. Healey and D. Slater, "Using Illumination Invariant Color Histogram Descriptors for Recognition", *Proc IEE Conf CVPR*, June 21-24, 1994, Seattle, WA, pp. 355-360.

[8] F.P. Incropera and D.P. DeWitt, *Fundamentals of Heat Transfer*, John Wiley and Sons, New York, 1981.

[9] G.J. Klinker, S.A. Shafer and T. Kanade, "Image Segmentation and Reflection Analysis through Color", *Proceedings of DARPA Image Understanding Workshop*, Cambridge, MA, 1988, pp 838 - 853.

[10] D.G. Lowe, "The Viewpoint Consistency Constraint," *International Journal of Computer Vision*, vol. 1, no. 1, 1987, pp. 57-72.

[11] J.D. Michel and N. Nandhakumar, "Unified Octree-Based Object Models for Multisensor Fusion", *Proc 2nd IEEE Workshop on CAD Model-Based Vision*, Champion, PA, 1994.

[12] N. Nandhakumar and J.K. Aggarwal, "Integrated Analysis of Thermal and Visual Images for Scene Interpretation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 4, July 1988, pp. 469-481.

[13] N. Nandhakumar and J.K. Aggarwal, "Thermal and Visual Information Fusion for Outdoor Scene Perception", *Proc. of IEEE International Conference on Robotics and Automation*, Philadelphia, PA, April 1988, pp. 1306-1308.

[14] N. Nandhakumar, "A Phenomenological Approach to Multisource Data Integration: Analyzing Infrared and Visible Data", *Proc. IAPR TC7 Workshop on Multisource Data Integration in Remote Sensing,* College Park, MD, June 14-15, 1990.

[15] N. Nandhakumar, "Robust Physics-based Sensor Fusion", Journal of the Optical Society of America, JOSA-A, 1994, to appear.

[16] S.K. Nayar and R.D. Bolle, "Reflectance Ratio: A Photometric Invariant for Object Recognition", *Proc IEEE ICCV*, 1993.

[17] T.H. Reiss, "Recognizing Planar Objects Using Invariant Image Features", *Lecture Notes in Computer Science, 676,* G. Goos and J. Hartmanis (Eds), Springer-Verlag, Berlin, 1993.

[18] E. Rivlin and I. Weiss, "Semi-Local Invariants", *Proc IEEE CVPR* 1993, pp. 697-698

[19] D. Weinshall, "Direct Computation of Qualitative 3D Shape and Motion Invariants", *IEEE Trans PAMI*, vol 13, no 12, Dec 1991.

[20] D. Weinshall, "Model-based Invariants for 3D Vision", *Proc IEEE CVPR* 1993, pp. 695-696

[21] I. Weiss, "Noise-Resistant Invariants of Curves", *IEEE Trans PAMI*, vol 15, no 9, July 1993, pp. 943-948

[22] L.B. Wolff, "Polarization-based material classification from specular reflection", *IEEE Trans PAMI*, Nov 1990, pp 1059-1071.

[23] M. Zerroug and R. Nevatia, "Quasi-Invariant Properties and 3D Shape Recovery of Non-Straight, Non-Constant Generalized Cylinders", *Proc IEEE CVPR 1993*, pp 96-103

[24] R. Zhang, P.-S. Tsai, J.E. Cryer, M. Shah, "Analysis of Shape from Shading Techniques", *Proc IEEE CVPR*, Seattle, WA, 1994, pp. 377-384.